The Run II Offline Computing Models
CD-Doc-2604-v1

Rob Kutschke

February 11, 2008

Abstract

This document describes the offline computing models used by CDF and D0 at the start of 2008. This is still a draft version.

Contents

1 The D0 Model

The D0 model is implemented as a set of Excel spreadsheets that I obtained from Amber Boehnlein in January. The xls files and the relevant worksheets in each file are given below:

1. `data_assumptions.xls`
   - `assumptions`, `data sizes`.
2. `hw_assumptions.xls`
   - `storage cost projections`, `fileserver projections`, `node infrastructure cost`, `CPU projections`
3. `processing_2008.xls`
   - `FNAL analysis costs`, `FNAL CPU costs`
4. `file_servers_2008.xls`
   - `tape costs`, `tape drives`, `Analysis Costs`.
5. `Global_Planning_2008.xls`
   - `total cost`

There are other worksheets in these files. Some contain notes that are useful for understanding the model but are not part of the model proper. Others contain computation of "value" which is a currency for accounting contributions made by outside institutions; the value calculations are outside of the scope of this project.

Appendix **??** contains printed versions of these worksheets, most with markup which is described in the text. It would be best to follow the discussion while browsing the worksheets with Excel.

1.1 Overview

In its broadest outline the D0 offline computing model is:

1. Jobs to be done at Fermilab.
   (a) All of the main data processing pass, including reco, fixing and skimming.
   (b) Main body of user analyses; in particular user analyses that require access to large datasets will be done here.
   (c) Provide the main data store, both archival (tape) and the disk pool, for all activities both onsite and offsite. Data sets produced offsite will be uploaded to FNAL for archival storage.
2. Jobs to be done offsite
   (a) Reprocessing, including reco, fixing and skimming. In the past some has been done a FNAL but that is not explicitly planned for in this model.
   (b) The main Monte Carlo production. All simulated events will be returned to FNAL for archival storage and for distribution to user analysis. In practice some Monte Carlo is run on site.
   (c) Additional user analysis will be done off site.

The model considers that Fermilab will need to deploy additional resources each year in order to meet the demands of the model. In particular this model projects the FNAL needs for:

1. CPU for the main processing pass and for user analysis.
2. Tape volumes in ENSTORE. For all data products.
3. Disk space for project space and for caching for accessing tape resident products.
4. File servers to serve the disk space.
5. Network.
6. Other infrastructure: racks, power, cooling.

The model foresees retiring CPU after 4 years. It also foresees retiring of tape volumes. *Just heavily used ones? Or...?*

1.2 The Model

The inputs to the model fall into the following classes:

- Properties of data:
  - Event rates, for both data and MC events.
  - Sizes of data products, per event.
  - Storage fractions on tape and disk.
  - Event processing times.
- Properties of Hardware
  - CPU power per node.
  - Cost of fileservers.

The model described in these spreadsheets starts in 2004, at which time 10^9 events had already been recorded by D0. This number of events, denoted by N_0 is given in cell C13 of worksheet `assumptions` in `data_assumptions.xls`. On a number of the worksheets, the formulae for 2004 are different than for subsequent years; *I presume that some of this is hack to give the right starting values for 2005? Because some quantities are cumulative across years it is necessary to start with an accurate description of the starting conditions. Is this right?*

1.2.1 Event Rates

The main body of the model is driven by the rate, averaged over 1 year, at which events are recorded from the experiment, row 8 of worksheet `assumptions` in `data_assumptions.xls`. Denote this by r_y. This number may change from year to year, hence the subscript y.

To be precise, this number is not actually an input to the model, it is computed from:

1. The peak rate at which events can be recorded, in Hz. This changes by year. Row 7 of worksheet `assumptions` in `data_assumptions.xls`. Denote this by R_{DAQ_y}.

2. A scale factor, a_y, that converts peak rate to the rate averaged over the year. This scale factor is also a function of the year and is the place to account for planned shutdowns. Row 4 of worksheet `assumptions` in `data_assumptions.xls`.

For further calculations, only row 8 of this worksheet is used, not rows 4 and 7.

A second input to the model is the rate, in Hz, at which D0Gstar MC events are to be generated, row 11 of worksheet `assumptions` in `data_assumptions.xls`. In some years this is specified as a fraction of R_y and in other years it is specified directly. Denote this by R_{MC_y}. For convenience cell C11 contains the rate in Hz that corresponds to 10^6 event s generated per week.

A third input to the model is the rate, in Hz, at which PMCS MC events are to be generated, row 12 of worksheet `assumptions` in `data_assumptions.xls`. This is always specified as a fraction of R_y. Denote this by R_{PMCS_y}.

Using the above notation one can define the corresponding per year event rates and the integral event rates. These calculations are done in lines 5 to 8 of the worksheet `data sizes` in `data_assumptions.xls`. The per year event rates are given by,

$$n_y = r_y \times 365 \times 24 \times 3600 \tag{1}$$
$$n_{MC_y} = r_{MC_y} \times 365 \times 24 \times 3600 \tag{2}$$
$$n_{PMCS_y} = r_{PMCS_y} \times 365 \times 24 \times 3600, \tag{3}$$

and the cumulative rate,

$$N_y = n_y + N_{y-1}. \tag{4}$$

The notation developed here is summarized in Table 1.

1.2.2 Sizes of Data Products

The full list of data products considered by the model is listed in cells B15:B31 of worksheet `assumptions` in `data_assumptions.xls`. These fall into four classes:

C0 The raw data and the output of the main processing pass, shaded yellow.

C1 The output of reprocessing, shaded green.

C2 The output of the D0Gstar chain, shaded magenta plus MC rootuple.

C3 The output of PMCS chain, also shaded magenta.

The reasoning behind the class structure will be described in the next section.

Column C of the worksheet gives the size per event of each data product, in MB. All of these sizes are direct inputs to the model.

Columns E through J of the worksheet define the yearly tape factor: a tape factor of 1 says that all events of this data product, produced in the specified year, will be written to tape; a factor less than 1 says that only a subset of the data will be written to tape. In the early years, or in years following major detector upgrades, it is assumed that some data products will be created several times; therefore tape factors greater than 1 are possible.

Columns L through Q give the corresponding disk factors for each data set. In this case factors above 1 are not present because it is assumed that superseded data products will be removed from disk.

In the following the subscript i denotes a data product, S_i denotes the size of a data product, in MB, t_{iy} denotes the tape factor for data product i and year y and d_{iy} denotes the disk factor for data product i and year y.

1. *What if the sum of disk + tape is less than 1. For example the TMB++ line in the green, re-reco section?*
2. *For raw data a disk factor of 0.1 is assumed for later years. Is this explicitly for data to be used by calibration?*

1.2.3 Rate of Production of Data Products

The distinction among the four classes of data products is the rate at which they are produced.

C0 Each year, all data recorded in that year will be passed through the main processing pass once. Normally this step will not be repeated in subsequent years.

C1 In 2004, no re-processing was included in the model. For years 2005 and forward, in each year, y, all data up to and including the data from year $y-1$ will be re-reconstructed. Data from the current year is not re-reconstructed.

C2 *I believe that the intention of the model is that D0Gstar derived data products will be produced each year at a rate given by r_{MC_y}. The spreadsheets are inconsistent in this aspect of the model. See section A.2.*

C3 *PMCS events are to be created at a rate proportional to r_{PMCS_y}. See section A.2.*

Within this information, one can define the yearly production rate for each data product, n_{yi},

$$n_{yi} = \begin{cases} n_y & \text{for i in class 0} \\ N_{y-1} & \text{for i in class 1} \\ n_{MC_y} & \text{for i in class 2} \\ n_{PMCS_y} & \text{for i in class 3} \end{cases} \tag{5}$$

For each data product the required tape and disk space each year is given by,

$$T_{yi} = \frac{n_{yi}}{1024^2} S_i \, t_{iy} \tag{6}$$

$$D_{yi} = \frac{n_{yi}}{1024^2} S_i \, d_{iy} \tag{7}$$

where the factor of 1024^2 converts from MB to TB. *See section A.1 for a typo in the worksheets.* The total disk and tape requirements, summed over all data products, in year y is:

$$T_y = \sum_i T_{yk} \tag{8}$$

$$D_y = \sum_i D_{yk} \tag{9}$$

The integral tape and disk requirements, in TB, through to year y,

$$I(T_y) = T_y + I(T_{y-1}) \tag{10}$$

$$I(D_y) = D_y + I(D_{y-1}) \tag{11}$$

The calculation in this section is done in worksheet `data sizes` of `data_assumptions.xls`.

Table 1 defines the notation used in this summary of the D0 model. Using this notation the derived quantities in the model are,

Table 1: Quantities used in the D0 Model. All the descriptions that say "up to" should be read as "up to and including".

Quantity	Unit	Definition
Event Rates		
R_{DAQ_y}	Hz	Peak DAQ event rate in year y.
a_y		Converts from peak event rate to average rate over the full year
r_y	Hz	Average DAQ event rate, averaged over a full year
r_{MC_y}	Hz	Production rate of D0Gstar (GEANT) events
r_{PMCS_y}	Hz	Production rate of PMCS events
n_y	y^{-1}	Yearly average data rate
n_{MC_y}	y^{-1}	Yearly production rate of D0Gstar (GEANT) events
n_{PMCS_y}	y^{-1}	Yearly production rate of PMCS events.
Data product sizes and storage factors		
S_i	MB	Size of data product i; not a function of time
t_{iy}		Fraction of data product i in year y to be tape resident
d_{iy}		Fraction of data product i in year y to be disk resident
n_{yk}		Number of events produced in year y of class k.
N_y		Number of events recorded up to and including year y
T_{yk}	TB	Tape required for data products of class k in year y
D_{yk}	TB	Disk required for data products of class k in year y
T_y	TB	Tape required for all data products in year y
D_y	TB	Disk required for all data products in year y
$I(T_y)$	TB	Integral of tape storage required through to year y
$I(D_y)$	TB	Integral of disk storage required through to year y

1.2.4 Projected CPU Capabilities and Costs

The worksheet `CPU projections` in `hw_assumptions.xls` contains the model of how CPU capabilities and costs will evolve with time. One number from this spreadsheet that is used later is D2, the cost per node. Denote this by $C_{\rm node}$; this should depend on the year but it is not used that way; the year by year cost information is in column J but that information is not propagated further into the model. *See Section A.6.*

Another number used later is D3, the IO cost per 100 nodes. Denote this by IO_{100}. What does this include? Routers, fibers, installation? Should this be year dependent?.

The node tax, D4, is not used in the model; instead the node tax is taken from cell C8 of the `node infrastructure` worksheet in `hw_assumptions`.

The main goal of this worksheet is to compute cells H8:H19, the cpu power of a single processor, measured in SpecInt's; this number is used in subsequent spreadsheets to compute the number of new nodes that are required each year. The cells bordered in red are the result of computations while all others are inputs to the model entered directly. The cells bordered in blue are commented on below.

Spreadsheet `processing_2008.xls`, which uses the numbers in cells H8:H19, assumes that a computing node contains a dual single core CPU; that is, it contains two CPU's each with the power listed in column H. For other CPU configurations, such as 2x2 cores or 1x4 cores, the number that should go into column H is half of the total power in the node. A handy number is that 1 GHz-s = 476 SpecInts.

The path from the left to the right across this spreadsheet is rather convoluted as different methods are used for different years. There are also some dead ends. *I presume that this reflects changes to the model with time. Right?* The shortcut to column H starts with column D, an expression of Moore's law for CPU power in SpecInt's, with a 2 year doubling time.

$$CPU(y) \quad = \quad CPU_0\, 2^{\frac{y-2000}{2}}, \tag{12}$$

where $CPU_0 = 335.58$ is the power of one CPU in the year 2000. Column H is computed from column D as given in Table 2. The factors of 2.6/2.8 and 2.6/3.2 are empirical. Cells H14 through H18 skip the model and just plug in numbers for recent and current equipment. The formula $476\times n\times 2$ describes the power in SpecInt's of a dual core x GHz processor. Recall that only half of the CPU power in the node should be specified in column H.

For completeness, the rest of the worksheet is now described. The "Nominal GHz" cells, B12:B17, are not used anywhere in the model. Cells D8 to D19 were described earlier. Cells E8:E19 scale the SpecInt value from D8:D19 to GHz. *I think there is bug here; see Section A.3.* Cells F12:F19 correct the numbers in column E based on local experience; this is the source of the empirical factors 2.6/2.8 and 2.6/3.2. Column G turns the numbers from F back into SpecInts; this step undoes the error made going from D to E. Column H was described previously. Column J is entered by hand.

Table 2: Details for computing column H in the `CPU Projections Worksheet`. Column D is just Moore's law with a two year doubling time and the numerical factors are empirical. The formula $476 \times x \times 2$ is described in the text.

Cell	Algorithm
H8	D8 $\times 2.6/2.8$
...	...
H13	D13 $\times 2.6/3.2$
H14	$476 \times 2.4 \times 2$
H15	H14
H16	$476 \times 4 \times 2$
H17	H16
H18	H16 $\times 1.25$
H19	D1 9$\times 2.6/3.2$

1.2.5 Projected Storage Capabilities and Costs

These worksheets are also found in `hw_assumptions.xls`.

The worksheet `storage cost projections` contains no computations; all cells contain inputs to the model that are entered directly.

The cells in worksheet `fileserver projections` mostly contain inputs to the model that that are entered directly. The exceptions are cells G13:G15.

G13 G12×(G12/G10); that is, geometric growth on a two year interval.

G14 G13; constant performance within a two year interval.

G15 This calculation is obsolete (per comment in J15). The value was based on fitting a curve to the time series of disk capacity data and extrapolating to 2011.

At present the full model is only integrated to 2009 so it is not a problem that there an obsolete entry in G15. I believe that C9 was used as scratch space or is part of a work in progress.

The cells in the worksheet `node infrastructure cost` are the start of a more sophisticated model of the care and feeding of nodes. All but C8 are inputs to the model entered directly; C8=SUM(C3:C7). Denote C8 by C_{infra}.

1.2.6 Resources Needed for the Main Processing

This computation is done on worksheet `FNAL farm costs` in `processing_2008.xls`. This worksheet is shown in Figure ???. In this figure, quantities entered by hand are highlighted with red borders while computed quantities are not highlighted. The blue highlighted boxes are discussed later.

Cells B10 through B17 give number of nodes acquired each year for the main data processing task; for B10:B14 these are actual acquisitions; for B15:B17 these are the projections of the model, taken from row 27. Denote this quantity by N_{node_y}, the number of nodes acquired in year y. *What was* *the actual number purchased for 2007?* Cells C:10:C17 give the computing power, in SpecInts, of the sum of the nodes represented in column B. Denote this by A_y, the power of the farm nodes acquired in year y. This computation uses the CPU power per processor computed in Section 1.2.4, column H in the `CPU projections` worksheet in `hw_assumptions.xls`. For year y, denote the CPU power per node by P_y, and the computation is:

$$A_y \quad = \quad 2 \times N_{node_y} \times P_y. \tag{13}$$

The factor of 2 comes from the assumption of two processors per node and the information in the `CPU projections` worksheet was computed thinking of this factor. *See Section A.5 for a possible bug in C17.*

Row 21 copies the event rate, in Hz, averaged over the year from row 9 of the `assumptions` worksheet in `data_assumptions.xls`. This was denoted by r_y in earlier sections. Row 22 specifies the efficiency of processing, ϵ_y. *I presume that this is the aggregate of things like jobs dying, running the wrong job etc. I presume that things like the startup/shutdown transients and IO waits are bundled into the time per event?* Row 23 gives the assumed contingency, c_y. Row 24 gives the reconstruction time in GHz-s, t_{Reco_y}. Row 25 gives the number of SpecInt's required to perform the main data processing, using the formula,

$$CPU_y \quad = \quad 476\,\frac{r_y\,t_{Reco_y}}{\epsilon_y}(1+c_y). \tag{14}$$

Row 26 describes the farm before the new nodes for the current year are added; denote this by F_y, the power of the farm in year y. This is computed by assuming a 4 year replacement cycle for nodes and assuming that only 80% of the nodes survive into their last year before replacement,

$$F_y \quad = \quad 0.8A_{y-3} + A_{y-2} + A_{y-1}. \tag{15}$$

From this one can compute the number of nodes that should be acquired in year y, A'_y, which is given in row 27,

$$A'_y \quad = \quad \frac{CPU_y - F_y}{2P_y} \tag{16}$$

where the factor of 2 has the same role that it did in Equation 13. The prime was added to the notation to distinguish the number of nodes that model says to purchase from the number of nodes that actually were purchased in previous years. *I don't know why row 27 displays as an integer? The formula does not force it to be an integer. And if you compute rows 28 or 30 from row 27 it's clear that the number still has a fractional part.*
Row 28 gives the cost, C_{node_y}, to purchase this many nodes. For 2005 and onward this is given by

$$C_{node_y} \quad = \quad A'_y\,C_{\rm node}, \tag{17}$$

where $C_{\rm node}$ was taken from cell D2 in the `CPU projections` worksheet in `hw_assumptions.xls`. There is an exception for year 2004; in that case the cost per node is taken from cell C5 in the present worksheet. *See Section A.6.* Row 29 computes the cost to supply IO for the newly acquired nodes. For 2005 onward this cost is given by

$$C_{IO} \quad = \quad \mathrm{INT}(A'^{y}/100) \times IO_{100}, \tag{18}$$

where INT denotes taking the integer part and where IO_{100} was defined in cell D3 of the `CPU projections` worksheet. An exception is the year 2004, for which the cost per 100 nodes is taken from cell C6 of the current worksheet. the *Is it right to round down? Probably this is a low level detail done at the end; it also depends on spare slots left over from previous years.* Row 29 computes the node tax for the year; the tax per node is denoted by C_{infra}, which was defined earlier. It is taken from cell C8 of the worksheet `node infrastructure cost`. This is also a node tax defined in cell D4 of the `CPU projections` worksheet but that number is not used. The total node tax is,

$$\mathrm{Tax}_{\mathrm{node}} \quad = \quad A'_y \, C_{infra}. \tag{19}$$

Actually this node tax computation is a dead end: the infrastructure costs in the final roll-up are put in by hand. Row 31 gives the number of nodes in the farm for the current year. For 2006 onward it is the sum of A_y, not A'_y, for the current year plus the previous 3. For years 2004 and 2005, it is the sum for the current year plus the previous 2. This does not take into account the expected death rate that was used in row 26. *Probably ok unless the death rate is much higher than assumed?*.

Rows 35 to 39 are not used for anything. B36:B39 are copies of quantities computed in row 52. C25:C29 are nonsense: they are the number of CPU's in 2002 (B10) multiplied by the CPU power per node that changes year by year.

1.2.7 Fixing and Skimming

1.2.8 Tape Costs

1.2.9 Tape Drives

1.2.10 Analysis Disk and Network

1.2.11 Global Roll-up

A Bugs and Questions

A.1 1028

On worksheet `data sizes` in `data_assumptions.xls`, rows 5 through 8. The factor to go from MB to TB is 1028^2 when it should be 1024^2.

A.2 MC Production Rates

About worksheet `data sizes` in `data_assumptions.xls`, rows 22, 23, 24. I claim that the rate factor for all of these should be X7, not X6. One possible subtlety is 2006 row 24 for which the rate factor is the previous years real data rate? Was this a special case or an error?

Also 2004 production is all proportional to B5 when it should be B7 or B8? This might be a hack that gives about the right number of events on tape and disk?

Same questions for rows 43 to 44. These are proportional to X5 but should be X7 or X8.

The net effect of this is that a small portion of the required disk and tape is over estimated by a large amount: the aggregate over estimate is on the scale of 5 to 10%.

A.3 Column E in `CPU projections`

In cells E8:E19 of worksheet **`CPU projections`** in **`hw_assumptions.xls`**, there is a bug. The intent is to scale the results of column D from SpecInt's to GHz-s. The problem is that the denominator is not the same for all rows, it is D10 for row 8 and D9 for all others. At various other places in the spreadsheets the conversion between GHz-s and SpecInts is:

$$1\ \mathrm{GHz} - \mathrm{s} \quad = \quad 476\ \mathrm{SpecInt}. \tag{20}$$

So I presume that the denominator should be D10 for all cells.

The error gets undone going from column F to G; however the error does screw up the empirical factors 2.6/2.8 and 2.6/3.2. These factors are used by hand in column H. Does that matter?

A.4 Column H in `CPU projections`

About the "old" in H6. Does this mean that the calculation is old or that these are "old SpecInt's"? Is there such a thing as old and new SpecInt's?

A.5 C17 in `FNAL farm costs`

This uses the power per node from 2008, not 2009. This is repeated in cell H27. Is this a typo or an intentional part of the model?

A.6 Row 28 in `FNAL farm costs`

In row 28 the cost per node comes from different places depending which column you are in. For 2004 it comes from C5 on this worksheet. For the other columns it comes from D2 on the **`CPU projections`** worksheet. A further problem is that this should be explicitly year dependent. As written one can only make it right for the year of interest. Or was this really an intended part of the model; that cost is approximately fixed and performance goes up?

B Notes

- Does the model consider that the fileservers may become bottlenecks?

- The acquisition model assumes 20% dead CPUs in the final year. Does this match reality?

C Printouts of the Worksheets

This section contains printouts of the worksheets. On some worksheets I have highlighted some cells with a colored outline border:

Red These are "odd man out cells"; that is, if most cells on a worksheet contain input data and only a few cells contain computed values, then the cells that contain the computed values are bordered in red; if, on the other hand, most cells contain computed values, then the cells bordered in red are those that contain input data. On some worksheets there are sufficiently few cells that this notation is not necessary.

Blue Highlighted because they are discussed in the text.

Green Highlighted because I have questions about the intended computation in these cells.

data _assumptions.xls

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	data assumptions																
2																	
3																	
4		peak to average		2.9				2.13	2.25	2.13							
5		peak to weekly		2													
6						2005	2006	2007	2008	2009							
7		peak event rate				75	100	100	100	100							
8	rates	average event rate	16	Hz		25.86207	34.48276	46.94836	44.44444	46.94836							
9		weekly average				37.5	50	50	50	50							
10		raw data rate	5	MB/s													
11		Geant MC rate	1.653439153	Hz		3.88	3.45	16.00	16.00	16.00		!10 M week average					
12		PMCS MC rate	0	Hz		3.88	3.45	4.69	4.44	4.69							
13		Events collected	1.00E+09														
14			size		tape factor 2004	tape factor 2005	tape factor 2006	tape factor 2007	tape factor 2008	tape factor 2009		disk factor 2004	disk factor 2005	disk factor 2006	disk factor 2007	disk factor 2008	disk factor 2009
15	sizes	raw event	0.2	MB	1	1	1	1	1	1	raw event	0.01	0.01	0.1	0.1	0.1	0.1
16		raw/RECO	0.5	MB	0	0	0	0	0	0	raw/RECO	0	0	0	0	0	0
17		data TMB++(reco)	0.15	MB	2	1	3	1	1	1	data DST	0	0	0	0	0	0
18		data tmb++(fixed)	0.15	MB	3	1	0	1	1	1	data TMB	1	0	0.1	0.1	0.1	0.3
19		data tmb++(skim)	0.15	MB	0	1	1	1	0.9	1	data tmb+(skiml	0	0	0.25	0.25	0.25	0.25
20		CAF	0.05	MB	0	1	1	1	0.9	1	CAF	0	0	0.75	0.75	0.9	0.75
21		user format	0.01	MB	0.5	0.5	0.5	0.5	0.5	0.5	user formats	1	0	0.5	0.5	0.5	0.5
22		data tmb++(rereco)	0.15	MB	0	1	0	0.1	0	0.5	data tmb++(rereco)	0	0	0	0	0	0
23		data tmb++fixed	0.15	MB	0	1	1	0.1	0.25	1	data tmb++fixed	0	0	0	0	0	0
24		data tmb++skim	0.15	MB	0	0.7	0.7	0.1	0.225	0.7	data tmb++skim	0	1	0	0	0	0
25		CAF	0.05	MB	0	0.7	0.7	0.1	0.225	0.7	CAF	0	1	0	0	0	0
26		user format	0.01	MB	0.5	0.5	0.5	0.1	0.01125	0.5	user formats	1	1	0	0	0	0
27		MC D0Gstar	0.7	MB	0.01	0.01	0.01	0.01	0.01	0.01	MC D0Gstar	0	0	0	0	0	0
28		MC D0Sim	0.3	MB	0	0	0	0	0	0	MC D0Sim	0	0	0	0	0	0
29		MC DST	0.3	MB	0.2	0	0	0	0	0	MC DST	0	0	0	0	0	0
30		MC CAF	0.05	MB	1	1	1	1	1	1	MC TMB	0.1	0.1	0.3	0.3	0.3	0.1
31		PMCS MC	0.02	MB	0	1	1	1	1	1	PMCS MC	0	0.5	0.5	0.5	0.5	0.5
32		MC rootuple	0.02	MB	0	0	0	0	0	0	MC rootuple	0	0	0	0	0	0
33																	
34	calendar assumptions																
35		min	60	s													
36		hour	60	min													
37		day	24	h													
38		day	86400	s													
39		year	365	d													
40		year	8760	h													
41		year	31536000	s													
42		year one	2003														
43																	
44	rate increase assumptions																
45		rate factor	2														
46		phase_1	2														
47		phase_2	4														
48		last year	2009														
49		total years	6														
50		raw size factor	1.25														
51		down year	2005														

worksheet: assumptions (detail)

Figure 1: Detail of worksheet `assumptions` from `data_assumptions.xls`. The yellow, green and magenta shaded regions, are discussed in the text. Most cells contain values that are inputs to the model but the cells with red borders red contain computed values. The cells with blue borders are discussed in the text.

data _assumptions.xls

	A	B	C	D	E	F	G
1	data samples (events)						
2		**Current**	**2005**	**2006**	**2007**	**2008**	**2009**
3							
4							
5	events collected	1.00E+09	8.16E+08	1.09E+09	1.48E+09	1.40E+09	1.48E+09
6	total events		1.82E+09	2.90E+09	4.38E+09	5.79E+09	7.27E+09
7	Geant events		1.22E+08	1.09E+08	5.05E+08	5.05E+08	5.05E+08
8	PMCS events		1.22E+08	1.09E+08	1.48E+08	1.40E+08	1.48E+08
9	TAPE data accumulation (TB)						
10	raw event	189.25	154.35	205.80	280.20	265.26	280.20
11	raw/reprocessing	0.00	0.00	0.00	0.00	0.00	0.00
12	data TMB+ (reco)	283.88	115.76	463.06	210.15	198.94	210.15
13	data TMB+(fix)	425.82	115.76	0.00	210.15	198.94	210.15
14	data TMB+(skim)	0.00	115.76	154.35	210.15	179.05	210.15
15	CAF	0.00	38.59	51.45	70.05	59.68	70.05
16	user format	4.73	3.86	5.15	7.01	6.63	7.01
17	data TMB+ (reco)	0.00	141.94	0.00	41.21	0.00	410.58
18	data TMB+(fix)	0.00	141.94	257.70	41.21	155.55	821.15
19	data TMB+(skim)	0.00	99.36	180.39	41.21	140.00	574.81
20	CAF	0.00	33.12	60.13	13.74	46.67	191.60
21	user format	0.00	4.73	8.59	2.75	0.47	27.37
22	MC D0Gstar	6.62	0.81	0.72	29.04	38.32	48.13
23	MC D0Sim	0.00	0.00	0.00	0.00	0.00	0.00
24	MC DST	56.78	0.00	0.00	0.00	0.00	0.00
25	MC TMB	47.31	5.79	5.15	23.87	23.87	23.87
26	PMCS MC	0.00	2.32	2.06	2.80	2.65	2.80
27	MC rootuple	0.00	0.00	0.00	0.00	0.00	0.00
28	**Yearly storage (TB)**	**1,014**	**974**	**1,395**	**1,184**	**1,316**	**3,088**
29	**total storage (TB)**	**1,014**	**1,988**	**3,383**	**4,567**	**5,883**	**8,971**
30	**MC Yearly (TB)**	**111**	**9**	**8**	**56**	**65**	**75**
31	**MC Total (TB)**	**111**	**120**	**128**	**183**	**248**	**323**
32	legacy		421.09	506.82	140.10	342.68	2025.51
33	new		544.09	879.81	987.71	908.51	987.71
34							
35	disk data accumulation (TB)						
36	raw event	1.89	1.54	20.58	28.02	26.53	28.02
37	raw/reprocessing	0.00	0.00	0.00	0.00	0.00	0.00
38	data DST	0.00	0.00	0.00	0.00	0.00	0.00
39	data TMB	141.94	0.00	15.44	21.02	19.89	63.05
40	data TMB+	0.00	0.00	38.59	52.54	49.74	52.54
41	CAF	0.00	0.00	38.59	52.54	59.68	52.54
42	user format	9.46	0.00	5.15	7.01	6.63	7.01
43	data TMB+ (reco)		0.00	0.00	0.00	0.00	0.00
44	data TMB+(fix)		0.00	0.00	0.00	0.00	0.00
45	data TMB+(skim)		141.94	0.00	0.00	0.00	0.00
46	CAF		47.31	0.00	0.00	0.00	0.00
47	user format		9.46	0.00	0.00	0.00	0.00
48	MC D0Gstar	0.00	0.00	0.00	0.00	0.00	0.00
49	MC D0Sim	0.00	0.00	0.00	0.00	0.00	0.00
50	MC DST	0.00	0.00	0.00	0.00	0.00	0.00
51	MC TMB	4.73	3.86	15.44	21.02	19.89	7.01
52	PMCS MC	0.00	7.72	10.29	14.01	13.26	14.01
53	MC rootuple	0.00	0.00	0.00	0.00		
54	**Yearly storage (TB)**	**158**	**200**	**118**	**161**	**162**	**203**
55	**MC**	**0**	**12**	**26**	**35**	**33**	**21**
56	**Yearly legacy storage (TB)**	**158**	**0**	**0**	**0**	**0**	**0**
57	**total storage (TB)**	**158**	**212**	**144**	**196**	**196**	**224**
58	**MC Yearly (TB)**	**5**	**12**	**26**	**35**	**33**	
59	**MC Total (TB)**	**5**	**16**	**42**	**77**	**110**	
60							
61			adjusted for new formats				

	H	I	J	K	L	M	N
19		Scale factors for first column are all wrong.					
20		Scale factors for mc are below. Top 3 rows should be X7.					
21		Lower right should be G7.					
22		B5	C7	D7	E6	F6	G6
23		B5	C5	D7	E6	F6	G6
24		B5	C6	C5	E6	F6	G6
25		B5	C7	D7	E7	F7	G7
26		B5	C8	D8	E8	F8	G8
27		B5	C7	D7	E7	F7	G8
48		Scale factors are all X5. Should be X7 or X8 as for tapes.					

data sizes

Figure 2: Worksheet `data_sizes` from `data_assumptions.xls`. The worksheet proper is columns A:G; the information to the right are notes that are discussed in the text. Almost every cell in this worksheet contains a computed value; the exception is B17:B21 which contain the constant 0. I believe that there are errors in the rows bordered in green; see Section A.2. The rows in blue were collapsed in the worksheet as received; there may be some errors in these rows too; these cells are not propagated further in this model.

hw _assumptions.xls

	A	B	C	D	E	F	G	I	J	K
1			2005	2006	2007	2008	2010			
2	LTO II tape cost($)		50	40	35	35	35			
3	LTO II tape capacity(GB)	200								
4										
5	9940b tape cost($)		80	80	80	80	80			
6	9940b tape capacity(gb)	200								
7										
8	LTOIII tape cost ($)			50	50	40	35			
9	LTOIII tape capacity(gb)	400								
10										
11	LTOIV tape cost($)					115				
12	LTOIV tape capacity(gb)	800								
13										
14	ADIC Slot cost	$8.50								
15	New STK	$50.00								
16										
17										
18		Tape Drive Cost Estimate								
19	cost/tape ($)		115		year	relative year	Capacity(TB)	drive rate (mbytes/sec)		
20	STK series ($)		30,000		2003	0	0.2	20		
21	LTO series ($)		8,000		2004	1	0.2	20		
22	Mover node ($)		3,500		2005	2	0.2	20		
23	LTO III($)		12,000		2006	3	0.2	40		
24	LTO IV($)		$15,000		2007	4	0.2	40		
25					2008	5	0.8	40		
26					2009	6	1.0	80		
27					2010	7	1.0	80		
28					2011	8	1.0	80		

storage cost projections

Figure 3: Worksheet `storage cost projections` from `hw_assumptions.xls`. All quantities on this worksheet are inputs to the model; there are no cells with computed values.

hw _assumptions.xls

	A	B	C	D	E	F	G	I	J
1									
2									
3									
4									
5		IDE File Server Cost Estimate							
6	cost/fileserver		30,000		year	relative year	Capacity(TB)		
7	Network cost/16 FS		10,000		2003	0	2.5		
8					2004	1	3.5		
9	06 cost		21750		2005	2	15.0		
10					2006	3	19.0		
11					2007	4	36.0		
12					2008	5	36.0		
13					2009	6	68.2		
14					2010	7	68.2		
15					2011	8	65.3		!obsolete

fileserver projections

Figure 4: Worksheet `fileserver projections` from `hw_assumptions.xls`. Most of the cells on contain inputs to the model; the exceptions are the cells bordered in red than contain computed values; G13 is obtained by scaling G12*(G12/G10); G14 is set to G13; and G15 is the result of an obsolete attempt to curve fit the capacity time series. So far the model is only integrated to 2009 so a bad value here is not an issue.

hw _assumptions.xls

	A	B	C
1			
2			
3		Cisco port	330
4		Wiring	30
5		Floor space	0
6		Sysadmin	0
7		Electricity	0
8			360

node infrastructure cost

Figure 5: Worksheet `node infrastructure cost` from `hw_assumptions.xls`. Most cells contain inputs to the model; the cell outlined in red is the SUM(C3:C7).

hw _assumptions.xls

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2		Cost/node:		2,300										
3		I/O Cost/100 nodes		25,000										
4		Tax/node		500										
5		CPU expectations												
6		Calculation courtesy Steve Timm						0ld						
7		"Nominal GHz"	Old	SpecInt	GHZ	GHZ, corrected for buying cycle		SpecInt, buying cycle adjusted	FY purchased	cost/node				
8	1999		Old	238	0.5			221						
9	2000		Old	336	1		312	312						
10	2001	1.1	Old	475	1.4		441	441						
11	2002		Old	672	2		624	624						
12	2003	2.6Hz	Old	950	2.8	2.6	882	882						
13	2004	3GHz	Old	1343	3.9	3.1	1091	1091			Adjusted 8/4/2004 for buying cycle			
14	2005	4GHz	Old	1899	5.6	4.9	1646	2284.8	2005	$2,800	Adjusted 5/18/2005 for buying cycle			
15	2006	6GHz	Old	2685	7.9	5.2	1747	2284.8	2006	$2,245	Adjusted 4/20/2006 for buying cycle			
16	2007	10GHz	Old	3797	11.3	9.1	3057	3808		$3,000	Adjusted for the quad cores			
17	2008	15GHz	Old	5370	15.9	12.9	4334	3808		$2,300				
18	2009		Old	7594	22.6	18.3	6148	4760						
19	2010		Old	10739	31.9	25.9	8702	8702						
20			Old											
21														
22														
23														
24														
25														
26		Note: Am pretending that the dual cores are two single cores 9/4/2005--should fix this after the Shank review.												
27		Note: 6/07--reminder: column "H" is a "per/processor" and the spreadsheet assumes two processors, thus the h16 corresponds to about 16 GHZ for the quad co												
28		Note: assuming that the cost goes down in 2008, but the processor remains as in 2007 (like with the 2005/2006 situation)												

CPU Projections

Figure 6: Worksheet `CPU projections` from `hw_assumptions.xls`. The cell outline in red is the SUM(C3:C7). The other cells are values input by hand.

processing_2008.xls

	A	B	C	D	E	F	G	H
1								
2								
3								
4								
5	Cost/node:		2,000					
6	I/O Cost/100 nodes		25,000					
7								
8								
9	2001	#of nodes	GHZ					
10	2002	260	324,480					
11	2003	96	169,344					
12	2004	120	261,840					
13	2005	160	731,136					
14	2006	200	913,920					
15	2007	157	1,195,895					
16	2008	21	161,277					
17	2009	101	767,693					
18		Primary Reconstruction Cost Estimate						
19								
20	Year		2004	2005	2006	2007	2008	2009
21	Average Rate		0	37.5	34.48275862	46.94835681	44.44444444	44.44444444
22	efficiency		70%	80%	80%	80%	80%	80%
23	contingency		20%	30%	20%	20%	20%	20%
24	Reco time		30	55	85	91	90	90
25	Required CPU			1595344	2092759	3050423	2856000	2856000
26	Existing system		0	552614	1128451	1854528	2694723	2088307
27	Nodes to purchase		160	228	211	157	21	101
28	Node Cost		$320,000	$524,833	$485,361	$361,155	$48,705	$231,840
29	Networking Cost		$25,000	$50,000	$50,000	$25,000	$0	$25,000
30	node tax			$82,148	$75,970	$56,529	$7,623	$50,400
31	#Nodes at FNAL		476	376	576	637	538	479
32								
33						note: 2007/8 assume replacement of		
34						2004/5 equipment		
35	2004	0	567,320					
36	2005	47	1,188,096					
37	2006	21	1,188,096					
38	2007	45	1,980,160					
39	2008	29	1,980,160					
40								
41		FIXING/skiming cost						
42								
43	Year				2006	2007	2008	
44	duration			90	90	90	90	
45	fraction			300%	100%	100%	100%	
46	Average Rate		0	314.6551724	139.8467433	190.4016693	180.2469136	
47	efficiency		70%	70%	70%	70%	70%	
48	contingency		0%	0%	0%	0%	0%	
49	Reco time		30	1	1	1	1	
50	Required CPU			213966	95096	129473	122568	
51	Existing system		0	0	0	213966	95096	
52	Nodes to purchase			47	21	45	29	
53	Cost		$0	$107,694	$47,864	$103,717	$65,734	

FNAL farm costs

Figure 7: Worksheet `FNAL farm costs` from `processing_2008.xls`.

processing_2008.xls

	A	B	C	D	E	F	G	H
1								
2								
3								
4								
5	Cost/node:		2,000					
6	I/O Cost/100 nodes		25,000					
7								
8								
9	2001	#of nodes	GHZ					
10	2002	160	199,680					
11	2003	200	352,800					
12	2004	120	261,840					
13	2005	120	548,352					
14	2006	160	731,136					
15	2007	96	732,041					
16	2008	133	1,010,326					
17	2009	455	4,334,014					
18		Analysis						
19								
20	Year		2004	2005	2006	2007	2008	2009
21	Average Rate		1405.42328	3.00E+03	4.80E+03	7.25E+03	9.57E+03	1.20E+04
22	efficiency		70%	70%	70%	70%	70%	70%
23	contingency		20%	20%	20%	20%	20%	20%
24	Reco time		0.5	0.45	0.4	0.3	0.3	0.5
25	Required CPU		573413	1102320	1566717	1774313	2341628	4901506
26	Existing system		552480	542069	764702	1042272	1331301	1434294
27	Nodes to purchase		10	123	176	96	133	455
28	Cost		$19,187	$281,989	$403,675	$221,073	$305,114	$1,047,083
29	networking			$25,000	$25,000	$0	$25,000	$100,000
30	#Nodes at FNAL		480	440	400	376	389	684
31								
32						note: 2007/8 assume replacement of		
33						2004/5 equipment		

FNAL analysis costs

Figure 8: Worksheet `FNAL analysis costs` from `processing_2008.xls`.

fileservers_2008.xls

	A	B	C	E	F	G	H	I
1								
2								
3								
4								
5								
6								
7								
8								
9	2004							
10								
11								
12								
13								
14	Contingency		0%					
15				2005	2006	2007	2008	2009
16		Data Volume		974	1,395	1,184	1,316	3,088
17		# to retire		0	0	0	0	10000
18		years volume		4871	6973	5918	1646	3089
19		replacements		0	0	0	2000	4167
20		purchase		4871	6973	5918	3646	7256
21								
22		Tape Cost		$ 243,550	$ 278,920	$ 207,130	$ 419,290	$ 253,960
23		cost to duplicate raw data		155	155	155	155	
24				775	775	775	194	
25				$ 89,125	$ 89,125	$ 89,125	$ 22,310	

tape costs

Figure 9: Worksheet `tape costs` from `fileservers_2008.xls`.

fileservers_2008.xls

	A	B	C	D	E	F	G
1							
2							
3							
4			Tape drives and mover nodes				
5							
6			2005	2006	2007	2008	2009
7							
8	#of drives		4	3	6	15	10
9	#of additional movers		3	3	3	0	0
10	total cost		10500	45000	79500	$277,500	155000

tape drives

Figure 10: Worksheet `tape drives` from `fileservers_2008.xls`.

fileservers_2008.xls

	A	B	C	D	E	F	G	H	I
1									
2									
3		File Server Cost Estimate							
4		2001	#of servers	TB					
5		2002	0	0					
6		2003	20	50					
7		2004	32	112					
8		2005	15	225					
9	2004	2006	19	361					
10		2007	13	468					
11		2008	12	432					
12		2009	9	614					
13									
14	Contingency		40%						
15				2004	2005	2006	2007	2008	2009
16		Data Volume (TB)		0	212	144	196	196	224
17		Project Volume			35	24	65	49	37
18		total volume		190	247	168	262	245	262
19		contingency		40%	40%	20%	50%	20%	40%
20		amount to replace		0	0	0	50	112	225
21		years volume (# servers)		18	24	11	13	12	9
22		replacements		0	0	0	0	0	0
23		#purchase		18	24	11	13	12	9
24		#owned		18					
25		Cost		$ 540,000	$ 720,000	$ 330,000	$ 390,000	$ 360,000	$ 270,000
26		Networking		$ 20,000	$ 20,000	$ 10,000	$ 10,000	$ 10,000	$ 10,000
27		total cost		$ 560,000	$ 740,000	$ 340,000	$ 400,000	$ 370,000	$ 280,000
28		total volume		0	387	748	1,166	1,486	1,875
29		equivelent file servers		0	37	48	49	50	33
30		value		$ -	$ 1,110,000	$ 1,440,000	$ 1,470,000	$ 1,500,000	$ 990,000
31		Networking value		$ -	$ 30,000	$ 30,000	$ 40,000	$ 40,000	$ 30,000
32		Total value		$ -	$ 1,140,000	$ 1,470,000	$ 1,510,000	$ 1,540,000	$ 1,020,000
33									

Analysis Costs

Figure 11: Worksheet `Analysis Costs` from `fileservers_2008.xls`.

Global Planning_2008.xls

	A	B	C	D	E	F	G	H	I	J
1										
2										
3										
4				Purchased 2003	Purchased 2004	Purchased 2005	Purchase 2006	Purchase 2007	Purchase 2008	Purchase 2009
5	FNAL Analysis CPU			$470,000	$277,000	$343,291	$453,628	$480,410	$305,114	$804,947
6	FNAL Reconstruction			$200,000	$370,000	$638,927	$545,423	$474,917	$48,705	$370,670
7	File Servers/disk			$111,000	$350,000	$400,000	$ 1,400,000	$1,150,000	$ 360,000	$975,000
8	Mass Storage			$280,000	$254,700	$19,600	$57,000	$97,500	$277,500	$175,000
9	Infrastructure			$244,000	$140,000	$347,020	$100,000	$100,000	$100,000	$100,000
10	FNAL Total			$1,305,000	$1,391,700	$1,748,839	$2,556,051	$2,302,828	$1,091,319	$2,425,618
11										
12										
13										
14										
15										
16										
17										
18										
19										

total cost

Figure 12: Worksheet `total cost` from `Global Planning_2008.xls`.